Recognition of Typed Letters in Noise

DONALD W. BROWN

Applied Research Laboratory, Sylvania Electronic Systems,* Waltham, Massachusetts

Ten human observers were required to guess the identity of typed letters which were presented in varying amounts of noise. Various criteria were applied to the data, and rank order correlations between criteria were obtained. The data are summarized in the form of a confusion matrix.

I. INTRODUCTION

Interest in letter recognition (which may be considered as a part of the more general problem of pattern recognition) dates back to the work of Roethlein (1912). Roethlein surveyed most of the available typefaces by presenting letters for identification at varying distances from the subject. By means of this procedure he was able to rank letters according to their legibility.

Tinker (1928) employing tachistoscopic presentation, produced legibility scales for printed letters (upper and lower case), numerals, and some mathematical signs.

Most recently, Neisser and Weene (1960) studied the legibility of hand-printed characters. Results were presented in the form of a confusion matrix of the type used in the present investigation.

The present study represents an attempt to determine a baseline for investigation of recognition of hand-printed characters by humans and machines. It seemed appropriate, as a first step, to determine the accuracy with which humans recognize printed characters presented in varying degrees of noise.

II. METHOD

APPARATUS

There were twenty-two carbon copies of the alphabet in capital letters, made on light copying paper. An IBM electric typewriter was

* A Division of Sylvania Electric Products Inc.
FIG. 1. Samples of the twenty-second, sixteenth, tenth, and fourth carbons of the stimuli, L, O, and H.

used to insure equal key pressure. The paper was then stapled on individual 2½ X 3 in. cards, one letter per card. The cards were arranged in order of decreasing “noise” or blur, making a pile with the original copy of the letter on the bottom.

Samples of the stimuli are presented in Fig. 1. The twenty-second, sixteenth, tenth, and fourth carbons are shown for the letters: L which was the easiest to recognize (i.e., had least number of errors); O which was intermediate in difficulty; and H which was most difficult to recognize (had the greatest number of errors).

SUBJECTS

The subjects were ten male employees of the Applied Research Laboratory (ARL), Sylvania Electronic Systems, a division of Sylvania Electric Products Inc.

PROCEDURE

Subjects (Ss) were instructed to begin with the twenty-second copy of a pile, picked at random, and proceed through to the first copy, giving a response or guess as to which letter was printed on each card. Each guess was also to be given a confidence rating by S, from one (1) to five (5): 1 indicating no confidence and 5 indicating complete confidence.
RECOGNITION OF TYPED LETTERS IN NOISE

TABLE I
THE CORRELATIONS OF THE SEVEN CRITERIA WITH ONE ANOTHER

<table>
<thead>
<tr>
<th></th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.81⁺</td>
<td>0.62⁺</td>
<td>0.91⁺</td>
<td>0.94⁺</td>
<td>0.58⁺</td>
<td>-0.24⁺</td>
</tr>
<tr>
<td>2</td>
<td>0.88⁺</td>
<td>0.81⁺</td>
<td>0.86⁺</td>
<td>0.42⁺</td>
<td>-0.34⁺</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>0.54⁺</td>
<td>0.66⁺</td>
<td>0.24⁺</td>
<td>-0.30⁺</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td></td>
<td>0.97⁺</td>
<td>0.66⁺</td>
<td>-0.20⁺</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>0.64⁺</td>
<td></td>
<td></td>
<td>-0.20⁺</td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>+0.13⁺</td>
<td></td>
</tr>
</tbody>
</table>

⁺ 1% level = 0.487.
⁺⁺ 5% level = 0.381.

The series of piles made up the whole alphabet, but S understood that any letter might be repeated at any time.

III. RESULTS

The letters were ranked according to seven criteria:
1. The trial of earliest correct recognition.
2. The first trial of continual correct recognition.
3. The trial with a rating of 5.
4. Cumulative total of correct ratings.
5. The mean correct rating (from 4).
6. The Roethlein (1912) rank order.
7. The Tinker (1928) rank order.

These rank orders were then correlated with one another (Table I).

The five alphabet ranks from this experiment all intercorrelate at the 1% level of significance. The Roethlein rank order correlates best with the weighted ranking criteria (4 and 5), and the Tinker rank order does not correlate significantly with any.

A confusion matrix for trails 4 through 22, with most frequent errors indicated, is presented in Table II. Entries indicate the number of times each response was made to each stimulus. The number of errors made on each stimulus as well as the most frequent error(s) appear in the lower margin; the number of times each letter occurred as an erroneous response, as well as the letters to which the response was given most frequently, appear in the right margin.

The present data was ordered according to the number of erroneous responses made to each stimulus character. The same was done for the
TABLE II
CONFUSION MATRIX FOR TRIALS 4–22 WITH TOTAL ERRORS AND MOST FREQUENT ERRORS

| STIMULUS | A 88 54 | B 78 32 | C 48 | D 12 | E 101 | F 123 | G 497 | H 41 | I 150 | J 308 | K 499 | L 267 | M 102 | N 897 | O 447 | P 106 | Q 447 | R 21 | S 3133 | T 19 | U 1653 | V 497 | W 52 | X 97 | Y 77 | Z 77 |
|-----------|---------|---------|------|------|-------|-------|-------|------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|
| STIMULUS  | A 88 54 | A 88 54 | A 88 54 | A 88 54 | A 88 54 | A 88 54 | A 88 54 | A 88 54 | A 88 54 | A 88 54 | A 88 54 | A 88 54 | A 88 54 | A 88 54 | A 88 54 | A 88 54 | A 88 54 | A 88 54 | A 88 54 | A 88 54 | A 88 54 | A 88 54 | A 88 54 | A 88 54 | A 88 54 | A 88 54 |
| RESPONSES | B 78 32 | B 78 32 | B 78 32 | B 78 32 | B 78 32 | B 78 32 | B 78 32 | B 78 32 | B 78 32 | B 78 32 | B 78 32 | B 78 32 | B 78 32 | B 78 32 | B 78 32 | B 78 32 | B 78 32 | B 78 32 | B 78 32 | B 78 32 | B 78 32 | B 78 32 | B 78 32 | B 78 32 | B 78 32 |
| MOST FREQUENT ERRORS | D 12 | D 12 | D 12 | D 12 | D 12 | D 12 | D 12 | D 12 | D 12 | D 12 | D 12 | D 12 | D 12 | D 12 | D 12 | D 12 | D 12 | D 12 | D 12 | D 12 | D 12 | D 12 | D 12 | D 12 | D 12 |

Neisser and Weene (1960) data using only the letters (i.e., the numerals were not considered either as stimuli or as erroneous responses to letters). The rank order correlation between these two sets of data was —0.08.

IV. DISCUSSION OF RESULTS

Table I shows negative rank order correlations between the data of Tinker (1928) and the data of the present study (regardless of the recognition criterion used). Significant positive correlations, however, were found between the Roethlein (1912) data and four of the five measures used in this study. These differences (between the Tinker on the one hand and Roethlein and the present study on the other) may be due to differ-
ences in procedure. While Tinker employed tachistoscopic method of presentation, Roethlein varied legibility by varying the distance of the stimulus from the subject and the present investigation used varying amounts of "noise" to vary legibility. Thus in Tinker's study the subject was presented with the complete character with all elements visible for a very short period of time, while in the Roethlein study and the present investigation, under the conditions of long viewing distance and high noise level the outlines of the letters were visible (i.e., their general shape) but the internal elements (e.g., the horizontal bar of H) were obscured (c.f. examples in Fig. 1).

The small negative correlation (−0.08) between the data of Neisser and Weene (1960) and those of the present study would appear to indicate that there is little or no relationship between recognition of blurred typewritten letters and the recognition of hand-printed letters presented with good figure-ground contrast. It should be noted, however, that it is possible that the correlation found between these sets of data could be due to the particular set of hand-printed characters, especially if there was any tendency on the part of the individuals doing the printing to be more careful with those letters which they felt were difficult to recognize. Thus, if sets of hand-printed letters were graded according to the extent to which they resembled printed letters, it would be expected that those most resembling printed letters would yield a high correlation while those bearing least resemblance would yield a low or even negative correlation.

Received: May 3, 1963

References


Tinker, M. A. (1928), The relative legibility of the letters, the digits, and of certain mathematical signs. J. Gen. Psychol. 1, 472–495.